International Journal of Basic and Applied Sciences, 14 (SI-3) (2025) 34-40

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/8pc16c12
Research paper

Development of Speech Recognition Technologies By Artificial in-Telligence Systems: A Review of Promising Research and Analysis of Current Achievements

Nodir Ramazonov, Nasiba Bozorova, Dildora Rakhmonova *, Nazokat Babadzhanova

Department of Uzbek Language and Literature, Tashkent State University of Law, Tashkent, Uzbekistan
**Corresponding author E-mail: drakhmonova8@gmail.com*

Received: May 6, 2025, Accepted: July 21, 2025, Published: august 15, 2025

Abstract

The relevance of this study lies in the widespread use of speech recognition technologies in everyday life and the vital role of artificial intelligence (AI) in this process. The main purpose of this paper is to review promising research in the development of speech recognition technologies by AI systems, as well as to investigate current achievements in this field. The methodological framework of this study is based on the methods of system analysis, which helped explore the most promising modern research in the development of AI for its subsequent application in speech recognition software. During this study, results were obtained indicating a significant number of promising domestic and foreign research studies on AI. The analysed studies also highlight the broad prospects for the use of AI-based speech recognition technologies for smart homes, voice search, the Internet of Things, and the Industrial Internet of Things, services for people with disabilities. The practical significance of the results obtained, as well as the conclusions of this paper, lies in their further application to create effective AI systems that can be used in various fields of modern science and technology for speech recognition.

Keywords: *Modern Technological Solutions; Natural Speech Processing; Voice Interface System; Linguistic Data; Internet of Things.*

1. Introduction

Artificial intelligence (AI) is a branch of computer science dedicated to developing systems that can execute activities usually necessitating human ability, including learning, reasoning, decision-making, and comprehending natural language. In speech recognition technologies, artificial intelligence is crucial. It enables computers to analyse and comprehend spoken language using methodologies such as deep learning, neural networks, and natural language processing. These systems transcribe audio input into text and provide actionable replies, serving as the foundation of contemporary voice-controlled apps and gadgets. Nonetheless, despite considerable advancements, numerous issues remain: the lack of linguistic resources for under-represented languages, difficulties in attaining high accuracy across various accents and dialects, apprehensions regarding data privacy and ethical usage, and constraints in emotional and contextual comprehension. Confronting these problems is crucial for the ongoing progress and wider acceptance of AI-driven speech recognition systems.

The advancement of speech recognition systems has progressed significantly over the last few decades, beginning in the 1950s with first efforts to identify isolated numbers and basic phrases using primitive acoustic models (Sharrab et al., 2025). During the 1970s and 1980s, research progressed with the introduction of Hidden Markov Models, allowing improved management of voice variability and enabling restricted vocabulary continuous speech recognition. The 1990s saw advancements in the integration of statistical methodologies and the creation of extensive annotated voice corpora, establishing a basis for commercial applications. The 2010s saw a significant turning point with the emergence of deep learning and neural networks, especially recurrent neural networks (RNNs) and subsequently transformer-based models, which substantially enhanced accuracy and real-time performance. Currently, AI-powered voice recognition systems are integrated into many applications, including virtual assistants, automated customer service, smart homes, and assistive technologies, signifying a transition from research prototypes to commonly used tools in daily life (Kiurchev et al., 2023; Kharchenko et al., 2017).

The systems that appeared in the 1970s now occupy a prominent place among the most promising innovative technologies. The advancement of this technology is largely due to both the emergence of the latest technological solutions and the development of modern, high-quality speech corpora and large computing power. The use of AI-based speech recognition systems in the languages of global communication, such as English, has become common, but the application for less-resourced languages, for example, Uzbek, remains a daunting task. Further research is needed to contextualise the application of AI-based speech recognition technologies in various segments of the Uzbek economy and assess the prospects of launching new AI-based speech recognition models or improving the existing ones.

AI-driven voice recognition systems are progressively used in Uzbekistan across several industries, notably in customer service, education, and technology (Porkodi, 2024; Azieva et al., 2021). AI-enabled personal assistants provide round-the-clock help and improve corporate productivity. In the education industry, AI technologies enhance learning experiences and reduce unemployment risks by equipping

graduates to engage with sophisticated systems (Shevchuk & Hunaza, 2025). Despite increasing use, obstacles persist, especially in under-resourced languages such as Uzbek, where the availability of linguistic datasets for training AI systems is limited.

Mukhamadiyev et al. (2023) investigated the use of AI personal assistants in various sectors of the economy. The researchers analysed the benefits of such assistants, as they can provide 24/7 support across a variety of sectors. Ongoing support not only increases the effectiveness and efficiency of a company's performance but also enhances its competitiveness by ensuring that customers' requests are instantly addressed (Destek et al., 2024). Considering existing benefits, AI-enabled assistants are seen as a key to a business's sustainability. The researchers further stressed that various sectors of the economy have already acknowledged the potential encapsulated in AI assistants and turned emerging technologies into an integral element of business performance. However, the researchers put it directly that the availability of AI-enabled technologies varies depending on the languages in which they are used. While technologies in English are gaining momentum, AI-enabled solutions in other languages, including Uzbek, are less developed; hence, they are not that common.

The topic of the applied AI systems was developed by Salaeva et al. (2023) in a scientific study of the role and importance of AI in speech recognition systems, especially for such less-resourced languages as Uzbek. The scientists concluded that for all the importance and usefulness of AI systems in speech recognition, humanity still has a long way to go to improve them. According to the authors, speech recognition systems are being introduced increasingly more: computers, smartphones, cars, and security systems are just a small area of AI applications. However, further progress in the Uzbek context is only possible with the creation of substantial linguistic data sets on which AI-enabled speech recognition systems can rely and the enhancement of the accuracy of these systems.

Bhable et al. (2023) conducted a systematic review of literature published between 2010 and 2020 and described the prospects and challenges of using AI-enabled speech recognition tools in the context of less-resourced languages, such as Uzbek. The researchers concluded that modern AI systems can recognise speech, give answers, and execute commands. Such systems have many advantages, including the presence of an elevated level of "humanity", self-learning, and automatic reproduction of actions. However, AI also has certain disadvantages; for example, these systems fail to provide an adequate transfer of data when it comes to scheduled and unscheduled dialects. Furthermore, there is a direct connection between the size of a linguistic data set and the accuracy of AI-based speech recognition technology. However, until relatively recently, little has been done to construct data sets that would encompass most used words and phrases in Uzbek.

Allah Rakha (2023, 2024) investigated the issues of the current state of AI in different sectors of the economy in Uzbekistan. It was concluded that AI systems should be considered as computer systems that use AI technologies in their operation. In Allah Rakha's opinion, the government's national AI strategy should focus on creating a conducive environment for AI systems, building a skilled AI workforce, and promoting the ethical use of AI. In addition, the regulatory framework should ensure that the development and deployment of AI technologies are in line with the country's laws and regulations. The same fully applies to AI systems used for speech recognition. Akmalov (2023) reached a similar conclusion. The researcher argued that Uzbekistan reports active use of AI-based technologies in different segments of the economy, including education.

The main purpose of this article is to review promising studies in the development of speech recognition technologies by AI systems, as well as to investigate current achievements in this field.

2. Materials and methods

The methodological approach of this study is based on the methods of system analysis. Some promising developments in the field of AI were reviewed for further use in the development of speech recognition software. The most relevant scientific achievements in this field were investigated. To complete the tasks set in this paper, a suitable theoretical base was formed. It included the results of an analysis of the most relevant scientific achievements in the development of AI speech recognition systems and their practical application in various fields of science and technology.

A systematic analysis of current scientific developments concerning AI allowed for the derivation of a definition of AI, as well as determining the scope of software and algorithms created based on AI systems. In addition, several domestic and foreign studies on scientific developments in the field of application of AI systems in modern science and technology were reviewed. Studies on the development of speech recognition software were also analysed.

An analytical study of the most relevant achievements in the field of application of AI systems for speech recognition allowed singling out both the achievements themselves and evaluating the factors that made these achievements possible. The use of this methodology determined the main stages of this study. The first stage defined the very concept of AI, as well as the stages of functioning of AI-based speech recognition software. Furthermore, the main problems of the functioning of such software related to the decoding of human speech by AI systems were identified. The main causes of such issues and workable solutions are briefly indicated.

The next stage of this study reviews promising modern scientific developments in the area under study. Such a review allowed comparing them, which, in turn, allowed for a value judgement concerning the role and significance of such technological solutions in various fields of modern science and technology. The final stage of this study analyzes the current advances in the use of AI systems for speech recognition. The most significant areas of practical application of AI systems in speech recognition were identified, and a brief analysis of current achievements in each of these areas was provided.

Furthermore, an analytical comparison was conducted between the obtained results and the findings and conclusions of other studies addressing the practical application of AI systems. This allowed refining the results obtained and forming the conclusions of this paper.

3. Results

AI is a broad concept that encompasses a wide scope of applications by computer software and algorithms, whose tasks include the implementation of technological solutions of various complexity. According to the definition of Allah Rakha (2023), AI is a science that studies and develops intelligent machines that can adequately respond to incoming external information while performing human functions. Mukhamadiyev and Cho (2024), designers of an AI-based speech recognition tool, UzAssistant, presented the main stages of the work of AI-based speech recognition software as follows:

1. Training the RNN-based models on the pre-processed dataset.
2. Applying hyperparameter tuning to optimise performance.
3. Validating the model by using a separate set of data.
4. Assessing the AI-based model's effectiveness and studying the opportunities for its generalisability.

The introduced algorithm of AI-based speech recognition software is universal and does not undergo significant changes when it is implemented in various types of speech recognition devices. A look at new research on AI-based speech recognition technologies wouldn't be complete without examining the scientific paper by Mavrovouniotis (1990), which explored how AI systems help create different levels of technology processes. That study noted the difficulties that arise in AI systems upon decoding human speech. These should include:

1. Difficulties in understanding the meaning of a single phrase when using various pronouns. In practice, such problems can be overcome by expanding the text base of AI-based software and adding new terms and phrases to it.
2. Speech recognition issues when using free word order and figures of speech in a sentence. This is solved by introducing models for constructing logical language correspondences into the software, which are inherent in a particular language of the recognised text.

Mavrovouniotis (1990) stressed that the described difficulties can only be resolved by introducing the above corrections into the computer algorithm, which are specific to the language in which the speech is recognised. Language and its expressive means are of key importance in the creation of specific AI-based computer algorithms that recognise speech (Bisenovna et al., 2024; Imamguluyev & Umarova, 2022).

Problematic aspects of speech recognition systems and AI are considered in the study by Marwala (2021). The author focused on the fact that AI systems are gradually filling the existing economic, social, and political space. In this context, AI systems that perform speech recognition are only a part of rational machines that are used in various decision-making processes. Furthermore, it is noted that when evaluating the performance of such systems, one should pay attention to the rationality and explainability of the results obtained. In other words, the level of efficiency of AI-based speech recognition systems is determined by how understandable, logical, and structured the text reproduced by such systems is. Without one of the above aspects, there can be no discussion on the rationality of the AI system's performance and its ability to qualitatively recognise speech.

The use of voice recognition technology has extensive social and economic effects (Bezshyyko et al., 2008; Kavaldzhieva, 2019). It alters human interactions with digital systems and revolutionises several areas of the labour market. These technologies enhance accessibility for individuals with disabilities, automate customer service, and improve efficiency in sectors like healthcare, education, and telecommunications. Conversely, they facilitate the automation of regular communication duties, perhaps eliminating jobs in sectors such as call centres and administrative assistance. This transition requires the enhancement of workers' skills and the restructuring of job responsibilities to conform to new AI-driven contexts. The common use of voice-enabled technologies changes how we communicate and raises questions about privacy, fair access, and including everyone, especially in areas with poor technology support (Babak et al., 2005). Consequently, whereas voice recognition technologies provide enhanced productivity and comfort, their use must be complemented by policies that address inequality and facilitate labour adaptation.

Relevant issues of speech recognition are raised in a scientific study by Mengliev et al. (2024). This study traces the dynamics of the development of AI systems for colloquial speech recognition, especially in a resource-constrained language environment. As explained by the authors, AI-enabled speech recognition is much less common in Uzbek than in other languages, which implies the need for future research in the field. The AI-enabled speech recognition model introduced by the researchers contains 1,600 sentences and 19,000-word forms annotated for parts of speech and named entities. The compiled linguistic dataset is expected to significantly increase the precision of AI-enabled speech recognition systems designed specifically for the Uzbek language. This precision is expected to compare with the precision demonstrated by AI speech recognition systems designed for other languages. Attempts to construct precise and customer-friendly speech recognition technologies are since AI-enabled speech recognition models are gaining momentum in the global market (Lyndyuk et al., 2024). The data collected by Statista reflects considerable growth in the AI-enabled speech recognition global market since 2020 and provides financial estimates of its further development (Fig. 1).

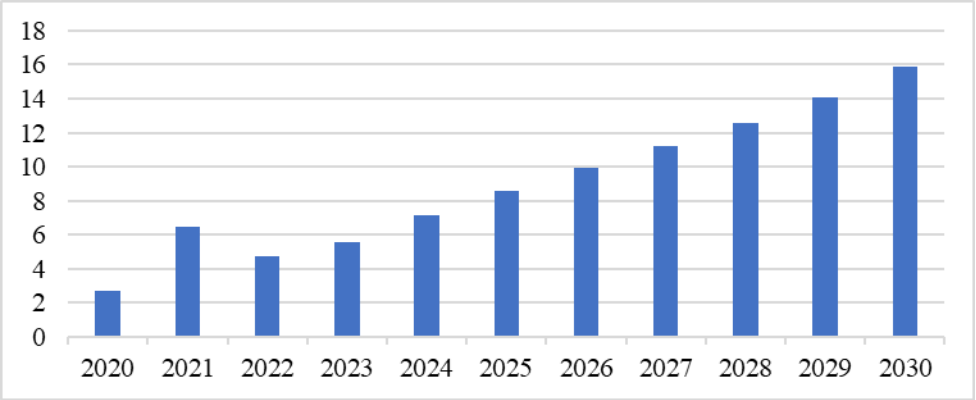

Fig. 1: AI-enabled speech recognition global market size, billion USD

Source: compiled by the authors based on data from Statista (2024).

The graph suggests that, despite some hurdles, the size of the global AI-enabled speech recognition market has increased two and a half times in four years: from 2.67 billion USD in 2020 to 7.14 billion USD in 2024. Furthermore, experts anticipate a doubling of the global market size, reaching 15.87 billion USD by 2030. Uzbek companies striving to get a share of this market are expected to accomplish a repertoire of tasks, including the most significant one, which is a growing precision of the AI-enabled speech recognition systems. Achieving sufficient levels of precision when transcribing conversations in Uzbek is expected to boost development across various sectors. AI speech recognition systems that rely on a significantly extended dataset "can be used to create applications such as improved chatbot systems, text mining applications, and other analytical tools for the Uzbek language, contributing to the development of those areas in the region for which those solutions were developed" (Mengliev et al., 2024).

Today, the communication of AI systems in a language that is close to human is far from flawless and requires further improvement (Trabelsi, 2024). The reason is that the speech of machines created using AI is almost devoid of the emotionality inherent in human speech, which significantly impoverishes speech in general, depriving it of expressiveness. The development of AI systems designed for speech recognition involves the introduction of new, more advanced algorithms that can solve more complex technological problems.

The use of AI in speech recognition is essential in smart house technology (Hassanien et al., 2022). Furthermore, such technologies are remarkably effective in voice search systems using the Internet of Things and Industrial Internet of Things technologies. Further development of technologies of speech recognition by AI systems will substantially affect the improvement of machine learning methods and

improve the quality of healthcare services, as well as services for people with disabilities. All the above actualises further research in this area and its importance.

The study by Potter & Oluwaseyi (2024) considered the issues of building robotic systems based on AI. Researchers note the possibility of using AI systems to develop mobile applications that process natural human speech. Apple mobile devices have already implemented similar technologies, enabling users to perform the following voice-activated functions:

– posting and commenting on social networks (Facebook, Instagram, Twitter).
– mobile device management, including function settings.
– building a route to move from one point to another, including the possibility of adjusting the route.
– voice search in Google, Bing, Yahoo.
– making outgoing calls and sending voice messages to instant messengers.

Today, scientific research continues in the field of practical application and improvement of AI technologies in the construction of robotic devices and the ultramodern technological systems (Potter & Oluwaseyi, 2024). The main challenges associated with using AI for robotics in speech recognition include:

1. Enhancing collaboration between humans and robots for achieving specific goals.
2. Achieving long-term autonomy, especially in less-resourced languages.
3. Scaling AI-powered robots for real-world situations, for example, education, and customer service.
4. Integrating AI-enabled models with other technologies, while not compromising real employees' needs.

A significant problem is the vulnerability of personal information obtained via voice conversations, which may result in privacy issues if improperly protected. AI systems that analyse voice data often depend on cloud storage, increasing the likelihood of unauthorised access or data breaches (Smailov et al., 2025b; Yessenbek et al., 2025). Moreover, mistakes in speech recognition, including the misreading of dialects or linguistic subtleties, may lead to unexpected data disclosure. The absence of strong legal frameworks in some areas intensifies these risks since AI systems may not consistently adhere to rigorous data protection requirements. Ensuring the safe and ethical management of speech data, while enhancing system accuracy, is essential for the broad use of these technologies.

To address the issues related to voice recognition systems, namely, with data privacy and security, many solutions may be used. Initially, strong encryption methods must be used for voice data during both transmission and storage, protecting sensitive information against unauthorised access. Integrating privacy-by-design principles in AI system development, such as limiting data collection and preserving only vital information, helps mitigate personal data exposure (Khan et al., 2025; Sasi et al., 2023). Moreover, implementing transparent data management techniques, such as clear user permission processes, could increase confidence in these platforms. Adherence to established data protection regulations must be prioritised to maintain stringent data privacy requirements. Ultimately, the ongoing enhancement of speech recognition models to effectively manage various dialects and diminish inaccuracies will mitigate the risks of misunderstanding and guarantee more dependable and secure systems.

International standards and regulations in voice recognition systems are crucial for assuring interoperability, user safety, data security, and ethical deployment. The International Organization for Standardization (ISO) (2017; 2022) established guidelines relevant to speech technologies, including ISO/IEC 30122-2 for voice user interfaces and ISO/IEC 27001 for information security management. Furthermore, international data protection legislation, particularly the EU's General Data Protection Regulation (2016), imposes stringent requirements for collecting, storing, and processing voice data, necessitating transparency, user permission, and data minimisation. In areas like North America, guidelines established by the National Institute of Standards and Technology (2025) provide criteria for assessing the precision and security of voice recognition systems. Compliance with these standards guarantees that technologies are both technically proficient and ethically responsible as well as legally compliant worldwide.

The ethical use of voice recognition technology presents considerable problems, especially pertaining to data privacy, security, and the protection of personal information (Khadzhiradieva et al., 2024; Kovalchuk, 2025). Given that these systems often gather and analyse sensitive voice data, there is an increased danger of unauthorised access, monitoring, and exploitation of personal conversations without the informed consent of users. Dependence on cloud-based infrastructures and ongoing data collection may increase the risk of breaches and identity theft if adequate protections are not established (Porkodi & Raman, 2025; Androshchuk, 2023). Furthermore, the lack of clear data management procedures and inadequate regulatory supervision may erode user confidence and violate data protection regulations. To address these ethical dilemmas, it is imperative to integrate privacy-by-design principles, guarantee user autonomy over data, and institute explicit accountability frameworks for developers and service providers.

In addition to the mentioned challenges, it is also critical to use AI robotic models to improve existing voice search systems. Voice search systems are used in modern search engines to speed up the input of a search query into the search bar (Smailov et al., 2025a; Denisov & Popovych, 2024). However, the clarity of the user's pronunciation is essential since, in case of any errors, the voice search system enters incorrect data, and the request cannot be fully completed. The voice query spoken by the user is instantly displayed in the browser's search bar, after which the search engine searches for information on the entered query. Thus, the time spent by the user on entering a search query from the keyboard of a desktop or mobile device is significantly reduced, and the search itself is accelerated. Today, the voice search function is actively used by users of search engines, gradually replacing the generally accepted way of entering a search query from the keyboard. Speech recognition in voice control systems is actively used in technologies such as "smart houses", in "smart" intercoms that open doors to residents when the flat number is pronounced, in the management of electronic household appliances, as well as in hotel rooms and restaurant businesses, equipped with voice control devices and intended for use by persons with disabilities. Furthermore, voice control systems are now actively used in the automotive industry for navigation devices, as well as in the aviation industry for controlling unmanned aerial vehicles.

Speech recognition technologies have considerable promise for education by improving learning accessibility, individualising training, and facilitating language development (Stepanenko, 2024; Dashko, 2023). These devices provide real-time transcription of lectures, enhancing notetaking and understanding for students with hearing impairments or learning disabilities. Speech recognition in language acquisition aids in pronunciation practice and conversational proficiency by delivering immediate feedback, thereby enhancing learner autonomy. Moreover, voice-activated educational tools and software provide dynamic learning experiences that cater to student requirements, enhancing engagement and efficiency. As education progressively incorporates digital platforms, voice recognition functions as a significant instrument in fostering inclusive, adaptable, and technology-enhanced learning settings.

Voice analytics of phone calls and conversations is used in business structures of various levels for quality assurance of interaction between employees and customers at all stages of conversation. This is done to improve the quality of service and increase sales, as well as to evaluate customer satisfaction with the quality of service and the feedback. Speech analytics performed by AI-based speech recognition systems allows raising business processes in a single organisation to a new level, identifying problems in the negotiation process and eliminating them promptly. The improvement of the negotiation process provides an increase in the conversion of telephone conversations

and increases the percentage of applications, which leads to an increase in the overall profit of the enterprise. Service systems with speech recognition are widely used in modern call centres and online banking systems, as well as in self-service systems, which make provision for the execution of specified operations after the input of voice commands. Systems of this kind can substantially reduce the time required to provide quality services to the client, which in the long term increases the total number of operations performed per unit of time, increasing the overall efficiency of the business process.

4. Discussion

Mengliev et al. (2024) considered several problematic aspects of the specifics of the use of AI technologies for speech recognition. As explained by the authors, the main problem with using AI-enabled speech recognition models and tools lies with a considerable shortage of neuro-linguistic programming resources for specific languages, including the Uzbek language. This shortage is connected to relatively modest "research and development in the field of computational linguistics in the Uzbek language". The researchers further argued that the detected deficiency can be partly addressed by considering linguistic theories and practices that have already been adopted for resource-rich languages, such as English. The suggested theoretical approaches include the linguistic theory of named entity recognition and part-of-speech tagging.

Muresan (2023) pointed to the effectiveness of AI-based speech recognition systems. The researcher stressed that the use of AI-enabled technologies in education can mitigate the risk for graduates to become unemployed because one-third of the workforce is being automated by the mid-2030s. As explained by the expert, this threat mainly concerns employees with a low level of education, while higher education graduates will remain in demand due to the ability to effectively cooperate with computers in creating "sophisticated systems that promote industry and wealth". Hence, the use of AI-enabled technologies in education can make graduates immune to the labour market's turbulence.

A nuanced analysis of using AI tools for educational purposes was also conducted by Tahir et al. (2024). The experts concluded that the application of AI tools in education is associated with such benefits as "greater flexibility in terms of time and space and a changing role for the tutor as a facilitator". With the rise of AI-enabled learning, education has become more student-focused, while learners have obtained greater control over the entire process. However, like Muresan et al. (2023), M. Tahir et al. considered not only benefits but also potential threats of using AI tools for learning. In addition to commonly mentioned data privacy concerns, the researchers also stressed the "lack of human touch" and equity and inclusion issues. The digital divide, which is particularly noticeable in rural and remote parts of the country, might deprive a considerable share of students of the benefits of revolutionised teaching and learning (Onipko & Yaprynets, 2024; Semenovska et al., 2023).

The application of AI systems in the field of machine learning was investigated by Pandey et al. (2022) in the study of AI and machine learning issues. The scientists concluded that the use of AI-based speech recognition systems allows achieving high quality in solving specific problems by converting analogue sound into digital impulses. The selected voice mode is stored on the hard disc and entered into memory at the time the program is launched (Syrlybayeva et al., 2024; Ahmadov, 2024). The effectiveness of speech recognition, according to the authors, depends on the technical characteristics of the hardware used. Similar conclusions were made during the present study, while in matters of the effectiveness of speech recognition, the proper use of computer technology involved in solving the tasks is also of significant importance.

Whetten & Kennington (2023) considered the practical application of automatic speech recognition (ASR) systems. The researchers compared the effectiveness of two common metrics, namely ASR and Word Error Rate (WER), in detecting errors in potentially sensitive conversations. To accomplish the task, the experts analysed five (N=5) conversations using different ASR systems (N=5). Based on collected data, the researchers concluded that the tested systems captured characteristics of ASR errors that cannot be captured by WER. Furthermore, the researchers emphasised the possibility of training an ASR system to incorporate severity as an element of the speech-transferring process. Therefore, the scientists proved that there is an ongoing improvement of existing speech recognition systems aimed at reducing the rate and severity of errors in transferred speech.

Nordin et al. (2023) investigated the key principles of the practical application of mobile AI. As a result, the researchers concluded that when using speech recognition software in mobile applications, one should pay attention to technical specifications, such as the amount of memory on the devices. The authors note that the program has its own dictionary of terms, the size of which directly depends on the amount of memory on the computer or other device where it is installed. Developers of speech recognition software should consider this fact when designing applications. Such a conclusion fully correlates with the results of the present study in the context of the specific features of the use of mobile devices for the technologies under study.

In a study of the fundamentals of applying AI to the development of modern technological systems, Hirsch-Kreinsen (2023) emphasised the fact that existing theoretical approaches fail to accurately reflect the dynamics of AI across various sectors. The researcher considered the technology-centred explanation of AI oversimplistic and examined the phenomenon through a social science lens, which describes AI-enabled solutions as a "promising technology". In the researcher's interpretation, the term means that successful integration of the novel technology into different segments can only be achieved by overcoming limitations and "previously unresolved challenges of AI diffusion and application". As explained by the expert, these limitations include the inability of AI-enabled systems to cope with "open worlds", incorporate crucial routine knowledge into the system process, resolve the explainable AI dilemma that has been in place since the 1990s, and estimate the amount of quality data needed for successful operation.

Nam et al. (2022) touched upon some problematic aspects of creating AI systems focused on the needs of an individual or a group of people. They concluded that when creating software for speech recognition using AI technologies, one should focus on the user interface, making it as understandable as possible for a potential user. The scientists note that interface issues may complicate the use of the AI-based speech recognition software, leading to many errors and accuracy issues. Kiran et al. (2023) stressed that despite arising challenges, the process of integrating AI speech recognition technologies into various sectors of the economy will persist. According to the experts, the integration process has become ongoing because of the considerable progress in ASR technology, which has significantly increased the precision of the data transfer process. Furthermore, the integration process is supported by the progress achieved in utilising deep learning models, such as RNN and transformer-based architecture. Multidisciplinary research to understand and address inherent limitations acts as a driving force behind the process of perfecting and integrating AI-based speech recognition solutions.

5. Conclusion

This study examined the most promising research and existing accomplishments in the use of AI for voice recognition technologies, highlighting both worldwide progress and the particular problems encountered in under-represented language situations, such as Uzbek. The study found that AI systems, particularly those that use deep learning and recurrent neural networks, have greatly increased the accuracy, adaptability, and usefulness of voice recognition technologies. Essential results emphasise the value of linguistic datasets, the advancement of language-specific models, and the need for enhanced management of intricate linguistic characteristics, including free word order and contextual ambiguity.

The research also revealed significant difficulties, including insufficient language resources, data privacy concerns, and ethical dilemmas, which persist in limiting the full potential of these technologies. Nonetheless, the practical implementations of AI-driven voice recognition, including smart homes, customer service, education, and assistive technologies, illustrate their increasing importance across several industries. Ongoing multidisciplinary research and adherence to international standards are crucial for addressing these difficulties and guaranteeing the responsible and inclusive development of speech recognition systems.

The current progress in AI speech recognition technology shows great potential for practical use and can help tackle many tasks related to operating different tech devices. Future research must prioritise the enhancement and diversification of linguistic databases, especially for under-represented languages like Uzbekistan, to improve the accuracy of voice recognition systems. Research on the integration of emotional expressiveness and contextual awareness in AI models is essential for enhancing natural interaction. Furthermore, research must focus on the elucidation and openness of AI decision-making to enhance user confidence and ensure regulatory adherence. Interdisciplinary research integrating linguistics, computer science, and ethics is crucial for addressing data privacy issues and guaranteeing equal access to voice recognition technology among diverse societal groups.

Acknowledgement

None.

References

[1] Ahmadov S 2024. Data encryption as a method of protecting personal data in a cloud environment. Bulletin of Cherkasy State Technological University, 29(3), 31–41. https://doi.org/10.62660/bcstu/3.2024.31

[2] Akmalov ZM 2023. The impact of artificial intelligence on the education system of Uzbekistan. Modern Scientific Research International Scientific Journal, 1(8), 193–201.

[3] Allah Rakha N 2023. Artificial intelligence strategy of Uzbekistan: Policy framework, preferences, and challenges. International Journal of Law and Policy, 1(1). https://doi.org/10.59022/ijlp.27

[4] Allah Rakha N 2024. Cybercrime and the law: Addressing the challenges of digital forensics in criminal investigations. Mexican Law Review, 16(2), 23–54. https://doi.org/10.22201/iij.24485306e.2024.2.18892

[5] Androshchuk A 2023. Cyber protection in programming languages Java and C#. Technologies and Engineering, 24(3), 9–14. https://doi.org/10.30857/2786-5371.2023.3.1

[6] Azieva G, Kerimkhulle S, Turusbekova U, Alimagambetova A & Niyazbekova S 2021. Analysis of access to the electricity transmission network using information technologies in some countries. E3S Web of Conferences, 258, 11003. https://doi.org/10.1051/e3sconf/202125811003

[7] Babak V, Filonenko S & Kalita V 2005. Acoustic emission under temperature tests of materials. Aviation, 9(4), 24–28. https://doi.org/10.1080/16487788.2005.9635914

[8] Bezshyyko O, Dolinskii A, Bezshyyko K, Kadenko I, Yermolenko R & Ziemann V 2008. PETAG01: A program for the direct simulation of a pellet target. Computer Physics Communications, 178(2), 144–155. https://doi.org/10.1016/j.cpc.2007.07.013

[9] Bhable SG, Deshmukh RR & Kayte CN 2023. Comparative analysis of automatic speech recognition techniques. In: Proceedings of the International Conference on Applications of Machine Intelligence and Data Analytics (pp. 897–904). Amsterdam: Atlantis Press. https://doi.org/10.2991/978-94-6463-136-4_79

[10] Bisenovna KA, Ashatuly SA, Beibutovna LZ, Yesilbayuly KS, Zagievna AA, Galymbekovna MZ & Oralkhanuly OB 2024. Improving the efficiency of food supplies for a trading company based on an artificial neural network. International Journal of Electrical and Computer Engineering, 14(4), 4407–4417. https://doi.org/10.11591/ijece.v14i4.pp4407-4417

[11] Dashko Y 2023. Information technologies in the educational process: Features and advantages. Pedagogical Sciences, 6(2), 46–51. https://doi.org/10.33989/2524-2474.2023.82.295094

[12] Denisov R & Popovych P 2024. Peculiarities of application of object recognation system in real time on microcontrollers with subsequent voise output of information for people with visual impairments. Technologies and Engineering, 25(3), 21–30. https://doi.org/10.30857/2786-5371.2024.3.2

[13] Destek MA, Hossain MR, Manga M & Destek G 2024. Can digital government reduce resource dependency? Evidence from the method of moments quantile technique. Resources Policy, 99, 105426. https://doi.org/10.1016/j.resourpol.2024.105426

[14] General Data Protection Regulation 2016. https://gdpr-info.eu/

[15] Hassanien AE, Chatterjee JM & Jain V 2022. Artificial intelligence and Industry 4.0. London: Academic Press.

[16] Hirsch-Kreinsen H 2023. Artificial intelligence: A “promising” technology. AI & Society, 39(4), 1641–1652. https://doi.org/10.1007/s00146-023-01629-w

[17] Imamguluyev R & Umarova N 2022. Application of Fuzzy Logic Apparatus to Solve the Problem of Spatial Selection in Architectural-Design Projects. Lecture Notes in Networks and Systems, 307, 842–848. https://doi.org/10.1007/978-3-030-85626-7_98

[18] International Organization for Standardization 2017. ISO/IEC 30122-2 – Information technology – User interfaces – Voice commands. https://www.iso.org/standard/63161.html

[19] International Organization for Standardization 2022. ISO/IEC 27001 – Information security, cybersecurity and privacy protection – Information security management systems – Requirements. https://www.iso.org/standard/27001

[20] Kavaldzhieva K 2019. The Impact of Digitalization on the Measurement of value in the production and operation of industrial products. In: 2019 International Conference on High Technology for Sustainable Development, HiTech 2019 (Article number: 9128260). Sofia: Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/HiTech48507.2019.9128260

[21] Khadzhiradieva S, Bezverkhniuk T, Nazarenko O, Bazyka S & Dotsenko T 2024. Personal data protection: Between human rights protection and national security. Social and Legal Studios, 7(3), 245–256. https://doi.org/10.32518/sals3.2024.245

[22] Khan MW, Destek MA & Khan Z 2025. Income Inequality and Artificial Intelligence: Globalization and age dependency for developed countries. Social Indicators Research, 176(3), 1207–1233.

[23] Kharchenko V, Ponochovnyi Y, Qahtan A-SM & Boyarchuk A 2017. Security and availability models for smart building automation systems. International Journal of Computing, 16(4), 194–202.

[24] Kiran B, Karthik V, Kumar Pavan P, Praveen Kumar K & Srinivasulu A 2023. Automatic speech recognition through artificial intelligence. International Journal of Multidisciplinary Research, 5(6), 1–10.
[25] Kiurchev S, Abdullo MA, Vlasenko T, Prasol S & Verkholantseva V 2023. Automated Control of the Gear Profile for the Gerotor Hydraulic Machine. In: Lecture Notes in Mechanical Engineering (pp. 32–43). Springer: Cham. https://doi.org/10.1007/978-3-031-16651-8_4
[26] Kovalchuk D. 2025. Utilising large language models for automated real-time cyber threat analysis. Bulletin of Cherkasy State Technological University, 30(1), 48–58. https://doi.org/10.62660/bcstu/1.2025.48
[27] Lyndyuk A, Havrylyuk I, Tomashevskii Yu, Khirivskyi R & Kohut M 2024. The impact of artificial intelligence on marketing communications: New business opportunities and challenges. Economics of Development, 23(4), 60–71. https://doi.org/10.57111/econ/4.2024.60
[28] Marwala T 2021. Rational machines and artificial intelligence. London: Academic Press. https://doi.org/10.1016/C2019-0-02529-0
[29] Mavrovouniotis ML 1990. Artificial intelligence in process engineering. Boston: Academic Press. https://doi.org/10.1016/B978-0-12-480575-0.X5001-1
[30] Mengliev D, Barakhnin V, Abdurakhmonova N & Eskulov M 2024. Developing named entity recognition algorithms for Uzbek: Dataset insights and implementation. Data in Brief, 54, 110413. https://doi.org/10.1016/j.dib.2024.110413
[31] Mukhamadiyev A & Cho J 2024. UzAssistant: Revolutionizing customer service in Uzbekistan with AI. Management and Future Technologies, 1(2), 12–16.
[32] Mukhamadiyev A, Mukhiddinov M, Khujayarov I, Ochilov M & Cho J 2023. Development of language models for continuous Uzbek speech recognition system. Sensors, 23(3), 1145. https://doi.org/10.3390/s23031145
[33] Mureșan M 2023. Impact of artificial intelligence on education. In: Balarabe K (Ed.), RAIS Conference Proceedings (pp. 81–85). Cambridge: Scientific Press. https://doi.org/10.5281/zenodo.8132828
[34] Nam C, Jung J-Y & Lee S 2022. Human-centered artificial intelligence: Research and Applications. London: Academic Press. https://doi.org/10.1016/C2020-0-02460-6
[35] National Institute of Standards and Technology 2025. Speaker Recognition. https://www.nist.gov/itl/iad/mig/speaker-recognition
[36] Nordin MN, Mustafa MZ & Mosbiran NF 2023. The application of artificial intelligence in android mobile learning for the special education students. In: 2023 International Conference on Artificial Intelligence and Smart Communication (AISC) (pp. 806–811). New York: Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/AISC56616.2023.10085531
[37] Onipko V & Yaprynets T 2024. Formation of readiness for innovative educational activities as a key competence of future specialists. Ukrainian Professional Education, 8(2), 71–78.
[38] Pandey R, Khatri SK, Singh NK & Verma P 2022. Artificial intelligence and machine learning for EDGE computing. London: Academic Press. https://doi.org/10.1016/C2020-0-01569-0
[39] Porkodi S 2024. The effectiveness of agile leadership in practice: A comprehensive meta-analysis of empirical studies on organizational outcomes. Journal of Entrepreneurship Management and Innovation, 20(2), 117–138. https://doi.org/10.7341/20242026
[40] Porkodi S & Raman AM 2025. Success of cloud computing adoption over an era in human resource management systems: a comprehensive meta-analytic literature review. Management Review Quarterly, 75(2), 1041–1075. https://doi.org/10.1007/s11301-023-00401-0
[41] Potter K & Oluwaseyi J 2024. Advances in artificial intelligence for robotics. Available at: https://www.researchgate.net/publication/377385099_Advances_in_Artificial_Intelligence_for_Robotics
[42] Salaeva M, Salaev U & Kuriyozov E 2023. Uzbek automatic speech recognition models using deep learning techniques. Computer Linguistics: Problems, Solutions, Prospects, 1(1), 218–223.
[43] Sasi S, Subbu SBV, Manoharan P & Abualigah L 2023. Design and implementation of secured file delivery protocol using enhanced elliptic curve cryptography for class I and class II transactions. Journal of Autonomous Intelligence, 6(3). https://doi.org/10.32629/jai.v6i3.740
[44] Semenovska L, Vazhenina I & Fazan V 2023. Individualization of learning as a development actuality information technological society. Pedagogical Sciences, 6(2), 30–34. https://doi.org/10.33989/2524-2474.2023.82.295073
[45] Sharrab YO, Attar H, Eljinini MAH, Al-Omary Y and Al-Momani WE 2025. Advancements in speech recognition: A systematic review of deep learning transformer models, trends, innovations, and future directions. IEEE Access, 13, 46925–46940. https://doi.org/10.1109/ACCESS.2025.3550855
[46] Shevchuk L & Hunaza L 2025. An analysis of international experience in implementing Artificial Intelligence in the educational process. Scientia et Societus, 4(1), 76–85. https://doi.org/10.69587/ss/1.2025.76
[47] Smailov N, Tsyporenko V, Ualiyev Z, Issova A, Dosbayev Z, Tashtay Y, Zhekambayeva M, Alimbekov T, Kadyrova R & Sabibolda A 2025a. Improving accuracy of the spectral-correlation direction finding and delay estimation using machine learning. Eastern European Journal of Enterprise Technologies, 2(5(134)), 15–24. https://doi.org/10.15587/1729-4061.2025.327021
[48] Smailov N, Uralova F, Kadyrova R, Magazov R & Sabibolda A 2025b. Optimization of machine learning methods for de-anonymization in social networks. Informatyka Automatyka Pomiary W Gospodarce I Ochronie Srodowiska, 15(1), 101–104. https://doi.org/10.35784/iapgos.7098
[49] Statista 2024. Speech recognition – Worldwide. Available at: https://www.statista.com/outlook/tmo/artificial-intelligence/computer-vision/speech-recognition/worldwide
[50] Stepanenko N 2024. Management of pedagogical communication: Theoretical and applied aspects. Ukrainian Professional Education, 8(2), 79–93.
[51] Syrlybayeva F, Kassymova X, Omarova E, Zhussipova B & Nurgalieva E 2024. Protection of information about employee's personal data in the Republic of Kazakhstan. Social and Legal Studios, 7(4), 90–102. https://doi.org/10.32518/sals4.2024.90
[52] Tahir M, Hassan FD & Shagoo MR 2024. Role of artificial intelligence in education: A conceptual review. World Journal of Advanced Research and Reviews, 22(1), 1469–1475. https://doi.org/10.30574/wjarr.2024.22.1.1217
[53] Trabelsi MA 2024. The impact of artificial intelligence on economic development. Journal of Electronic Business & Digital Economics, 3(2), 142–155. https://doi.org/10.1108/JEBDE-10-2023-0022
[54] Whetten R & Kennigton C 2023. Evaluating and improving automatic speech recognition using severity. In: Demner-Fushman D, Ananiadou S & Cohen K (Eds.), The 22nd Workshop on Biomedical Natural Language Processing and BioNLP Shared Tasks (pp. 79–91). Toronto: Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.bionlp-1.6
[55] Yessenbek Z, Markabayeva G, Albatyr I, Tleubayeva N & Atay S 2025. Application of Artificial Intelligence Technologies in Digital PR. Rotura Journal of Communication Culture and Arts, 5(1), 83–95. https://doi.org/10.34623/edqp-b256